ISSN 2347-1921

Bayesian Rule of Multiple Hypothesis testing: Past and Present

Muntazim Abbas Hashmi
COMSAT Institute of Information Technology,
Computer Science Department, Ada Peer Muraad, Vehari, Pakistan
muntazimabbas@gmail.com

Abstract.

The era of emerging mathematical techniques towards Hypothesis Testing is between 1908 to 2003. History of hypothesis testing is based on some distinct scientists which are discussed in this article. Several of works are there from different decades and from different authors, some of them have been considered briefly. Also works considered on two main different approaches and their comparison. The generalized concepts using different approaches towards hypothesis testing is discussed and compared with each other. Most impotent consideration in this article is the Bayesian decision theoretic approach and its generalization in the work of K. J. Kachiashvilli (2003). This work can be considered as a description of the problem related to the problems of multiple hypothesis testing and decision theory. To evaluate efficacy in multiple endpoints in confirmatory clinical trials is a challenging problem in multiple hypotheses testing which can be solved using this approach.

Keywords: Testing of Hypothesis; Bayesian Approach; Generalization of Bayesian Approach; Wald Sequential Analysis.

Council for Innovative Research

Peer Review Research Publishing System

Journal: Journal of Advances in Mathematics

Vol 2, No 1
editor@cirworld.com
www.cirworld.com, member.cirworld.com

ISSN 2347-1921

1. Testing of Hypotheses.

The basic results of statistical hypotheses testing were obtained by Wald [Wald, (1947, 1950)]. The detailed study on the fundamentals of this theory is contained in the monographs [Blackwell and Girshick, (1954); Lehmann, (1986)]. In works [Wald, (1947, 1950); Blackwell and Girshick, (1954); Jeffreys, (1961); Berger, (1985a); De Groot, (1973); Dickey, (1977)], special attention is given to the Bayesian criterion. The basic postulates of the classical theory are given in [Wilks, (1962); Rao, (2006); Kendall and Stuart, (1970); Stuart, Ord and Arnold, (1999); De Groot, (1970); Cramer, (1946); Zaks, (1971); Aivazjan, Yenyukov and Meshalkin, (1985) and others].

Depending on the chosen criterion, there are different classical methods of hypotheses testing [Blackwell and Girshick, (1954); Lehmann, (1986); Wilks, (1962); Rao, (2006); Cramer, (1946); Kendall and Stuart, (1970); Stuart, Ord and Arnold, (1999); Zacks, (1971); Berger, (1985 a, b)]: Neyman- Pearson's criterion, the Bayesian criterion, the maximum of posterior probabilities, the maximum likelihood criterion, Wald's sequential analysis and others. A lot of works are dedicated to the synthesis of optimal decision rules in theoretical and applied statistics [Casella and Wells, (1990); Kiefer, (1977); Robinson, (1979 a, b); Boratynska and Drozdowicz, (1999); Shi and Wan, (1999); Bagui and Datta, (1998); Liang, (1999); Samaniego and Vestrup, (1999); Westfall, (1997); Cheng, Su and Berry, (2003); Kachiashvili, (1989, 2003, 2006) and others].

To explore the properties of the papulation under the collected information can be divided into two parts, *estimation* and *hypothesis testing*. To test some assertion about some parameters of papulation is known as *hypotheses testing*. The most basic and general framework of the problem was given by Fisher (1890-1962),Abraham Wald (1902－1950), J. Neyman (1894-1981), and E. S. Pearson (1895-1980).

The story of hypothesis testing was formally originated from the introduction of *t test* in 1908, by Sir Ronald Aylmer Fisher (1890－1962) in his book *Statistical Methods for Research Workers (1925)*. He also did great number of works in this direction. The concept of testing a hypothesis on the basis of a test staistics using distribution of that statistics is called *test of significance*. This was firstly introduced by Fisher.

Most famous works of Wald are on *Statistical Decision Theory*, *Hypothesis Testing*, *Sequential Analysis* etc. Some of his important and basic works are discussed in this article.

From 1928 to 1934, Neyman and Pearson published seven of their ten most important papers on the theory of testing statistical hypotheses. In 1930 to 1933, they formed the likelihood ratio to find critical region for simple and composite hypotheses. From 1934 to 1972, Neyman made an outstanding work towards the theory of hypothesis testing. Neyman explained Wald's work following the same experimental designs. They introduced the problem of finding optimal areas solving a constrained optimization problem with the help of Lagrange Multiplier method. However there are other scientists who did an unfathomable work in this direction for example D. R. Cox, Maurice Stevenson Bartlett (1910－2002), Debabrata Basu (1924－2001), Allan Birnbaum (1923－1976), Alexander Philip Dawid (1946), Morris Herman DeGroot (1931－1989), William Edwards Deming (1900－1993) etc.

In 1959, E. L. Lehmann publish a book "Testing Statistical Hypotheses". This work was one of a great contributions towards this direction. This book covers several of topics on "Small Sample Theory" and "Large Sample Theory". He criticized, in some of his works, the approaches of Fisher and Neyman-Pearson. E. L Lehmann is also considered a big name in the theory of hypothesis testing.

In this work we will discuss some approaches, methods and would point out the deficiencies of the previous works. We will discuss some generalized approaches and try to reduce the emerging problems towards testing of hypothesis. Also we will discuss decision theoretic aspect of hypothesis testing.

There are two most important and effective approaches towards testing of hypothesis, *Frequentist* and *Bayesian*, associated with the interpretation of the concept of probability. Early works on testing of hypothesis are based on frequentist approach and did by, for example, Fisher, Cox, Jackson, Sawyer etc. Initially, the works did for the comparison of null and an alternative, which may be simple or composite. However, this takes no longer to make the people familiar with the concept of Multiple Hypotheses Testing. Both the approaches are used to solve this problem.

2. Multiple Hypotheses Testing

Several methods are used to interpret the best statements among some number of statements. Naturally, we prefer one which uses more information and simply gives the best decision. The decisions involving uncertainness must need some probabilistic concepts corresponding to inhale the incorrect decisions, in some sense. The methods which extend the logic and gives better reasoning about the uncertain statements involves concepts of probability.

Multiple Hypotheses Testing is just to choose one statement from a set of statements, called *family*, in the light of information by taking care of the errors made during this choice of hypothesis. This leads to us to make *inference* in such situations in which comparison is not required. Approximately all the situations deals with the controlling the family wise error rate. In the beginning, the problem was considered in frequentist point of view, *p*-values. Then is used the Bonferroni methods, which are extended and adjusted later. Then the methods related to Bayesian approach are used and improved till now. Most of the works in practical situations are dealing with these methods and trying to overcome the problems related to this. Cournot (1843) asserts about this concept of multiple testing but he admitted the problems related with it are not solved up to the optimality criteria. Later, after 1940, there are several works on this problem as Mosteller and Nair

(1948), Tukey (1949), Duncan (1951), Lehmann (1957), Miller (1981), A. L. Rukhin (1982), H. Wang (2003), J. P. Romano and M. Wolf (2005) etc [5][7]. These works are important due to there worth and applications.

Several books were published till 1994, however, the first book on multiple testing was given by Miller (1966). There are other books which are very important in this direction, for example, Multiple Comparisons (Klockars and Sax 1986), Multiple Comparisons for Researchers (Toothaker 1991), Multiple Comparisons Procedures (Toothaker 1993) etc [8].

3. Frequentist Approach.

Association of this idea is with the "frequency interpretation of probability", which is originated in different forms from the beginning of 19th century. However, between 1908 to 1970, a huge amount of work have been devoted to hypothesis testing using this approach. Several of mathematicians and statisticians uses this theory and derive many important results towards the direction of hypothesis testing. In the last two sections we have discussed some of the renowned names. In the next section we will, precisely, discuss some works using this approach.

4. Works on Frequentist.

- Abraham Wald (1939); *Contributions to the Theory of Statistical Estimation and Testing Hypotheses*, The Annals of Mathematical Statistics, Vol. 10, No. 4 (Dec., 1939), pp. 299-326

This was the first formal work on hypothesis testing. In this work, Wald discuss the theory of Neyman-Pearson, associated with errors of two types, and introduced the method of choosing best region of acceptance using some weight function and prior probabilities. He also discussed the deficiencies of Neyman-Pearson approach. He constructed risk function as expected value of loss function and find those areas in which the risk is minimum. He did not discuss the effects of the errors of two types separately. However, this work was a leading key towards the generalized conditional case discussed in the last topic of this paper.

- Abraham Wald (1941); *Asymptotically Most Powerful Tests of Statistical* , The Annals of Mathematical Statistics, Vol. 12, No. 1 , pp. 1-19.

By following some assumptions, Wald introduced some inequalities to obtained the regions which are asymptotically the best critical regions for hypothesis testing. He uses likelihood estimator and find that the distribution of it, is asymptotically normal. He shows that it is the powerful way of obtaining these regions.

- Abraham Wald (1943); *Tests of Statistical Hypothesis Concerning Several Parameters when the number of Observation is Large*, The Annals of Mathematical Statistics, Vol. 10, No. 4 , pp. 299-326

The work is associated with the problem of testing hypothesis about the parametrical values of a distribution on the bases of independent observations. Precisely, this was the most formal way to discuss on multivariate distributions. This paper contains several of results which are also very helpful in the modern research, for example, distribution of likelihood estimator is asymptotically multivariate normal, general problem can be reduced into multivariate normal, tests of simple and composite hypothesis with power of the tests, general composite hypothesis and some optimum properties of likelihood estimator etc.

- Abraham Wald (1943); *On the Efficient Design of Statistical Investigations*, The Annals of Mathematical Statistics, Vol. 14, No. 2, pp. 134-140

The problem of linear hypothesis is discussed in this work. A measure of the efficiency of the design of a statistical investigation for testing a linear hypothesis is also considered. Some designs are also given to check the effectiveness and efficiency of the method.

- Abraham Wald (1945); *Sequential Tests of Statistical Hypotheses*, The Annals of Mathematical Statistics, Vol. 16, No. 2, pp. 117-186

This work is concerned with the problem of sequential test of a statistical hypothesis and introduced a new procedure, the *sequential probability ratio test*. This work considers how to choose as critical region controlling the type I and type II errors. Wald divided the space into three different parts and defined the test as the observation results falls in a region, which decision will appear. He also considered the problem of computing the number of observations required by the sequential probability ratio test. This required small number of observations as compared to the other tests but more powerful. A merit of this procedure is, we do not need to determine the probability distributions in all the cases. Here is also discussed the problem of multiple comparison when even the hypothesis are not simple.

- D. R. Cox (1961); *Test of Separate families of Hypothesis*, Birkbeck College, University of london

This paper considers the problem of testing multiple hypothesis using maximum likelihood. He discussed a test based on this likelihood function. In this work,the problems of lognormal verses exponential are focussed.

- O. A. Y. Jackson (1968); *Some results on tests of Separate families of Hypotheses*; Biometrika, 55,2,p.355.

In the work of D. R. Cox, adequacy of the asymptotic results and the power function of the test is not considered. This work gives extension to the Cox's work and introduces the power function and discuss some other distributions in this respect. He also compares the results obtained from this work to the others.

- J. Kiefer (1977); *Conditional confidence Statements and Confidence Estimators*; Journal of the American Statistical Association, Vol. 72, No. 360, pp. 789-808.

He introduced a statistic, which he used to measure the confidence in interval estimation, hypothesis testing, selection problem in nonparametric and sequential procedures. He also compared the properties of the procedures with the others by using distribution of the introduced statistic.

- A. L. Rukhin (1982); *Adaptive Procedures in Multiple Decision Problems and Hypothesis Testing*, The Annals of Statistics, Vol. 10, No. 4, pp. 1148-1162.

This work is concerned with simple multiple decision problem and hypothesis testing. It is proved that the adaptive procedure estimates behave same as for a large number of observation as the minimax estimator behaves. He also gives some examples to show that there are situations in which adaptive procedures does not exists.

- K. R. Sawyer (1983); *Testing separate families of Hypothesis: An Information Criterion*; J. R. Statist. Soc. B 45,No. 1, pp. 89-99.

An information criteria is used to test the hypothesis and the bases of this work is obtained from the work of Cox. He measure the discrimination between the hypotheses using log-likelihood ratio and on the basis of obtained ratio, he made assertion about the acceptance or rejection of some hypotheses.

- K. R. Sawyer (1984); *Multiple Hypothesis Testing*; J. R. Statist. Soc. B 46,No. 3, pp. 419-424.

Sawyer discussed the method of comparison between a family of hypothesis. In this work, is introduced a test statistic which is just an extension of the work of Cox (1961). He showed that this statistic can be used to solve the problem of multivariate case.

- E. L. Lehmann (1993); *The Fisher, Neyman-Pearson Theories of Testing Hypotheses: One Theory or Two?*, Journal of the American Statistical Association, Vol. 88, No. 424, pp. 1242-1249

This is one of the important works of E. L. Lehmann. In this work, he compared the approach of Fisher with that of Neyman-Pearson. He pointed out the several visible distinctions of these two approaches e.g. *fixed level and p-values*, *ignoring the errors of two type*. He concludes, in the end, both the approaches are originated from the same bases and, in philosophical point of view, they are same. This work is a leading act towards considering the problem of multiple comparison.

- C. Goutis, G. Casella, M. T. Wells (1996); *Assessing Evidence in Multiple Hypotheses*, Journal of the American Statistical Association, Vol. 91, No. 435, pp. 1268-1277.

In this work, he construct a method for the multiple hypotheses testing by using some informative measure, measure of evidence, in multivariate case. He obtained this measure as a function of *p-values* for the considered hypothesis.

- Y. Benjamini, W. Liu (1999); *A step-down multiple hypotheses testing procedure that controls the false discovery rate under independence*, Journal of the Statistical Planning and Inference, 82, pp. 163-170 .

This paper considers the problem of testing multiple hypothesis by controlling false discovery rate which is more powerful than familywise error rate.They introduced a new procedure to test the hypothesis by introducing a new statistics, based on p -values. Also discussed the power of step-up (SUFDR) or step-down (SDFDR) multiple comparisons, and recommended that SDFDR better to work when number of hypothesis, $m \le 16$, otherwise SUFDR works better.

5. Bayesian Approach.

In 18^{th} century, Thomas Bayes (1702-1761) introduced the notion which leads to the direction of updating information about the assumption made about the unknown parameters of the distribution after obtaining data. He did not formalized the concept as much as it required, but, a French mathematician Pierre-Simon Laplace (1749-1827) in 1774, formalized the notion given by Thomas Bayes. In the 20^{th} century, the ideas of Laplace were further developed in two different directions, *objective* and *subjective* Bayesian. A *point of view* which is based on the statement of a person, but not on some collected information for the statement, which may be accepted or rejected is called *subjective view*. For example, if a man asked that he has headache, then we are able to accept or reject his statement. We are not able to tell how much he is talking correct about his problem. However, if he say that he had not slept since last twenty four hours and still working on some book and he has a headache, then we are quite sure about the truth of the statement and can say something, which may not be exact, how much pain he is feeling now. A *point of view* which is not only based on the statement of a person but also on the belief and experienced information of a person is called *objective view* [3][6][9][11][13][14].

After the inventions of different methods of hypothesis using frequentist view, the journey starts towards the testing of hypothesis using bayesian and decision theoretic point of view. Several of the names occur during 1960 to 2003, some are, Wald (1939), I J Good (1950),Lindley (1965), Partt (1965),De Finetti (1974), Novick (1981), Rubin (1987), Berry (1996), Louis (2001), Gill (2002), Congdon (2003) etc, writers of books and other informative material towards this direction.

Some important books on Bayesian approach are *Bayesian Data Analysis* by Andrew Gelman etc al (2004), *Statistical Decision Theory and Bayesian Analysis* by James O. Berger (1985), *Bayesian Methods for Data Analysis* by B P Carlin etc al (2004), *Introduction to Bayesian Statistics* by William M. Bolstad (2007), *Bayesian Computation With R* by Jim Albert (2009), *Bayesian Theory (Wiley Series in Probability and Statistics)* by José M. Bernardo, *Kendall's Advance Theory of Statistics volume 2B Bayesian Inference* by Anthony O' Hagan.

6 Works on Bayesian.

6.1 Bayesian Inference

The basic approach towards Bayesian inference to encapsulate *information* about the parameter of the *distribution* involved in the light of observed data and the *prior* knowledge. This information may be infer about *estimation* or about *hypothesis* etc. To obtained this posterior information, data and prior knowledge is used in Bayes theorem, but not all methods which use Bayes theorem are considered the part of Bayesian inference. Only those which uses this posterior information to make inference [12].

- D. B. Duncan (1965); *A Bayesian Approach to Multiple Comparisons*, Technometrics, Vol. 7, No. 2 .

He discussed the different approaches towards multiple comparison and introduced a bayes criteria to obtained a bayes rule by minimizing the bayes risk. He showed that this criteria leads us to obtained such an areas which are optimum in the sense that it controls the errors of two type. He formulate the problem in the sense that it always accept or reject one of the given hypothesis at some given level, however, several cases are not discussed in this work which naturally exits e.g. none of the hypothesis is true, two or more of them are true etc on the basis of observation results.

- R. Fortus (1979); *Approximations to Bayesian Sequential Test of Composite Hypotheses*, Technometrics, Vol. 7, No. 3, pp. 579-591.

The problem of approximation of Bayesian sequential test for composite hypothesis is considered. The procedure depends on the sample size n and the cost of sampling c. The test statistic used is obtained from the log likelihood function, and we show that the optimal Bayesian stopping rule may be approximated by a stopping rule which depends only on n, c, and two likelihood ratio test statistics.

- J. C. Naylor, et. al. (1982); *Application of a method for the efficient Computation of Posterior Distributions*, Journal of Royal Statistical Society. Series C, Vol. 31, No. 3, pp. 214-225.

This work approximates the complicated likelihood functions, statistical procedures based on posterior distributions, or integrated likelihood by Gaussian quadrature which leads to efficient calculation of posterior densities for a rather wide range of problems. This work is a good example of considering the problems of computations in statistical methods.

- B. Toman (1996); *Bayesian Experimental Design for Multiple Hypothesis Testing*, Journal of American Statistical Association, Vol. 91, No. 433, pp. 185-190.

Here is designed the experiment for k -decision problem and uses the expected loss function, Bayes risk function, which is the leading key for obtaining an optimal design. Several new optimality criteria based on Bayes risk have been introduced specifically for the problem of multiple hypothesis tests.

- Robert Weiss (1997); *Bayesian Sample Size Calculations for Hypothesis Testing*, Journal of the Royal Statistical Society. Series D (The Statistician), Vol. 46, No. 2, Special Issue: Sample Size Determination, pp. 185-191

The work considers the problem of computing sample size to make the bayes factor greater then a cut off of some prescribed size. If b_{01} denotes the Bayes factor in favor of H_0 against H_1, then sample size n is computed by using some pre-specified level k by minimizing type I and type II errors at some extent. In classical approach, for example, the sample size is calculated as

$$n_c = \frac{\sigma^2(z_\alpha + z_\beta)^2}{\mu_1^2}$$

where σ^2 is known variance of normal distribution, μ_1 is hypothetical value for unknown μ mean of the normal to be asserted, z_α, z_β are the values obtained from $\Phi(z_\alpha) = 1-\alpha, \Phi(z_\beta) = 1-\beta$, where $\Phi(x) = \int_{-\infty}^{x} exp(-\frac{1}{2}t^2)dt$.

Bayesian, methodology is to have the posterior $1-\alpha$ a probability content interval have width w. This methodology might select n to be the smallest integer greater than $(4w^{-2}z_{\alpha/2}^2 - \tau_1^{-1})\sigma^2$, where τ_1 is hypothetical variance under H_1.

- H. Peter, et. al. (1997); *A Bayesian perspective on the Bonferroni Adjustment*, Biometrika, Vol. 84, No. 2, pp. 419-427.

Work considers the problem of comparison between p-values and posterior distribution for multiple hypothesis testing. He considered the Bonferroni multiple testing problem which uses the adjusted p -values, p -values multiplied by number of hypothesis. The similar approach is used for adjusted posterior probabilities and relation between adjusted posterior and p -values given on some conditions that number of hypothesis is large enough and bayes factor is small enough.

- J. M. Bernardo, et. al. (2002); *Bayesian Hypothesis Testing:A Reference Approach*, Internationa Statistical Review, Vol. 70, No. 3, pp. 351-372.

Using the intrinsic discrepancy between two fully specified probability models which is the minimum expected log-likelihood ratio is considered as loss function and Bayesian criteria is used. The loss function is an information distance between the probabilities and related to the kullback information distance. This distance $d(\theta_0, x)$, which is expectation of intrinsic discrepancy, is used to make assertion about the null hypothesis as if $d > d^*$ we reject H_0. He discuss the Lindley's paradox, Rao's paradox, reference priors, and uses the the approach to explain the inconsistencies of univariate and multivariate normal frequentist hypothesis testing.

Some of other important works towards this direction are given below. These works are impotent in the sense that they gives new directions towards testing of hypothesis. There are also discussed the cases where we need to emphasize.

- V. E. Johnson (2004); *A Bayesian χ^2 test for Goodness-of-fit*, The Annals of Statistics, Vol. 32, No. 6, pp. 2361-2348.
- A. Zellner (2006); *Some aspects of the history of Bayesian information processing*, Journal of Econometrics, 138, pp. 388-404.

6.2 Bayesian Decision Theory

Instead of the data and prior information, there is another information which is called loss information, estimation of distraction after a decision, which can be consider to make decisions. Making assertion about the values of the parameters by taking care of all the three information leads us to the concept of decision theory. Identifying the values of some parameters, uncertainties and other issues relevant in a given decision, rationality of the decision, and the resulting optimal decision is the summary of the decision theory. Using the concepts of hypothesis testing, the next step is towards the formalization of the problems related to statistical decision theory. However, there is a good amount of material towards this problem using frequentist approach, but these are not much formal, informative and accurate as much as the Bayesian one, in many of the cases. Last 80 years are very important with respect to the research in the decision theory. The reason behind to connect the decision problem under uncertainty with statistics is to formalized the problems related to decision theory in such a way that we obtain some generalized formulation which can be applicable in many of the cases. The Bayesian approach is mostly used in two ways, to compare posteriors and using risk function. Modern Approaches towards this theory uses the concept of risk function. Some works related to the Bayesian decision theoretic approach are considered.[10].

- S. Weerahandi, etc al (1981); *Multi-Bayesian Statistical Decision Theory*, Journal of the Royal Statistical Society. Series A (General), Vol. 144, No. 1, pp.85-93 .

This work extends the work of Nash J. F. (1981) on bargaining problem. He introduced a utility function using prior information and explain that this method is applicable in hypothesis testing. He states that using the utility function for some two action problem, we can minimize the Nash product and uses randomized decision rule to make decision about the parameter value. He uses the gain instead of loss and do not use the explicit comparison of utilities. This work does not provides any information measure for how much we are sure about the truth of a particular hypothesis.

- S. Weerahandi and J. V. Zidek (1983); *Elements of Multi-Bayesian Decision Theory*, The Annals of Statistics, Vol. 11, No. 4 , pp. 1032-1046 .

Among the bunch of decisions, say $\mathcal{B}$, this work discuss the problem of choosing one, a randomized decision β using some decision function $\delta \in \mathcal{D}$ in two different ways, considering some sub-sampling space $\mathcal{B}' \subset \mathcal{B}$ or $\mathcal{B}$ as a sub-sample from a super-papulation $\mathfrak{B}$. The worth of δ depends upon posterior expected gain of utility. In this work is said that if Mahalanobis distance between hypothesis is small, we can consider the average of the means of corresponding posteriors, say $B(\cdot \mid \beta)$, of δ. The core of this is as we maximize over $\mathcal{D}$ and minimize over $\mathcal{B}$ to $\{B(\delta \mid \beta) - c_{sa}(\beta)\}$, where $c_{sa} = max\{B(\delta' \mid \beta) : \delta' \in \mathcal{D}\}$.

Several other works related to the problem of choosing losses for a particular problem, controlling the risk and other issues are discussed in detailed. Some of them are mentioned below.

- Charles Peter J. Kempthorne (1988); *Controlling Risks under Different Loss Functions: The Compromise Decision Problem*, The Annals of Statistics, Vol. 16, No. 4, pp. 1594-1608 .

• Charles Lewis, Dorothy T. Thayer (2009); *Bayesian Decision Theory for Multiple Comparisons*, Optimality: The Third Erich L. Lehmann Symposium, Institute of Mathematical Statistics.

7 Comparison between Bayesian and Frequentist.

Instead of the philosophical flaws of classical inference, the Bayesian inference considers parameters as random variable and the posterior distribution of the parameter is the key towards making inference of any type about the parameter. In the classical cases, the parameters are fixed quantities. There are several works on this topic, but here we discuss some of them.

Other differences are bayes use more information then the frequentist, parameters which are not of interest (*nuisance parameters*) effects in frequentist for making assertion about the others which are of interest but not in the case of Bayesian, repeated observation are required to make inference in frequentist but not in Bayesian, Bayesian make direct inference about the parameters of interest but frequentist makes about the statistic computed by the data even several of the observations are collected. These works that are discussed below mainly considers such type of differences.

• D. J. Bartholomew (1965); *Comparison of Some Bayesian and Frequentist Inferences*, Biometrika, Vol. 52, No. 1/2, pp. 19-35

• G. Zech (2002); *Frequentist and Bayesian confidence intervals*, Springer-Verlag 2002, EPJdirect C12, 1–81 (2002), DOI 10.1007/s1010502c0012

• Sarat C. Dass and James O. Berger (2003); *Unified Conditional Frequentist and Bayesian Testing of Composite Hypotheses*, Journal of Statistics, Vol. 30, No. 1, pp. 193-210

• D.A.S. Fraser, N. Reid (2003); *Strong matching of frequentist and Bayesian parametric inference*, Journal of Statistical Planning and Inference 103 (2002) 263–285

8 Generalization of Bayesian Rule of Multiple Hypothesis testing.

However, despite a variety of works dedicated to the problem of statistical hypotheses testing and, in particular, to the Bayesian criterion, there is no work where the problem considered below has been solved in the offered manner.

This part of the article is basically concerned with discussing a introducing a new method of Multiple Hypothesis Testing which concerned only with the problem when the hypothesis are simple and precise introduced by K. J. Kachiashvilli and its applications using normal distributions is done by Muntazim Abbas Hashmi and Abdul Mueed. However, for the the case of composite hypothesis testing and if the hypothesis are imprecise, the situation is very much difficult. This method is extended as a constrained optimization problem also.

To choose one between several assumptions is the basic core of all these works which we have discussed earlier and till now people think that only this type of problems can occur. There is another possibility that none of the given hypothesis is true and it might also be possible that two or more can be true. There was no method solving this problem in general, i.e. on the basis of observation results, none of the given hypothesis or two or more hypothesis can be true under the required criteria of testing and controlling both type of errors at the given extent. So we have to find such a method which can be applicable in these situations. This problem, in its general form, is considered by K. J. Kachiashvilli in 2003.

When we make a decision, we have a threat of obtaining some loss. This loss will be case of two types of mistakes, rejection of a true hypothesis (Type I error) or acceptance of a false (Type II error). Some times in practical situations, to reject a true will give more loss then to accept a wrong one and vice visa. So, according to the situation we have to control the errors of these two types when we are going to minimize the risk corresponding to a decision. The method introduced in this article is based on this concept. This is considered that the risk is based on two components, and minimize the Bayesian risk, means conditioning one component of risk and minimizing the other. This work is a good description of the possible cases of hypothesis testing and decision making. Kachiashvili K. J. considered the problem of choosing a Bayesian decision as a constrained and unconstrained optimization problem, originated two different directions. Here is also discussed the problem of obtaining a sample size for an optimum solution. Several other issues related with the calculations are also discussed and solved. The approach use in this article is closely related to the approach of Wald and that of Neyman-Pearson [1][2][8].

However, despite a variety of works dedicated to the problem of statistical hypotheses testing and, in particular, to the Bayesian criterion, there is no work where the problem considered below has been solved in the offered manner.

Let's consider $n-$dimentional random observation vector $x^T = (x_1,\dots,x_n)$ with probability distribution density $p(x,\theta) = p(x_1,\dots,x_n;\theta_1,\dots,\theta_m)$, given on $\sigma-$algebra of Borellian sets of space $\mathbb{R}^n$ $(x\in\mathbb{R}^n)$, which is called the sample space. By $\theta^T = (\theta_1,\dots,\theta_m)$ is designated the vector of parameters of distribution. In general, $n \neq m$. Let in m -dimensional parametrical space Θ^m be given S possible values of considered parameters $\theta^{i^T} = (\theta_1^i,\dots,\theta_m^i), i = 1,\dots,s, i.e., \theta^i \in \Theta^m; \forall i : i = 1,\dots,S.$ On the basis of $x^T = (x_1,\dots,x_n),$ it is necessary to make the decision namely by which distribution $p(x,\theta^i), i = 1,\dots,s,$ the sample x was born.

Let's introduce designations: $H_i : \theta = \theta^i$, is the hypothesis that the sample $x^T = (x_1, \dots, x_n)$ was born by the the distribution

$$p(x, \theta^i) = P(x_1, \cdots, x_n; \theta_1^i, \cdots, \theta_m^i) \equiv p(x/H_i), i = 1, \cdots, S;$$

$p(x/H_i)$ is the priori probability of hypothesis H_i; $D = \{d\}$-a set of solutions, where $d = \{d_1, \cdots, d_S\}$,it being so that

$$d_i = \begin{cases} 1, & if \ hypothesis \ H_i \ is \ accepted, \\ 0, & otherwise. \end{cases}$$

$\delta(x) = \{\delta_1(x), \delta_2(x), \dots, \delta_s(x)\}$ is the decision function that associates each observation vector x with a certain decision

$$x \xrightarrow{\delta(x)} d \in D$$

Γ_i is the acceptance area of hypothesis H_i , i.e. $\Gamma_j = \{x : \delta_j(x) = 1\}$. It is obvious that $\delta(x)$ is completely determined by the Γ_i regions, i.e. $\delta(x) = \{\Gamma_1, \cdots, \Gamma_S\}$. Let's introduce loss function $L(H_i, \delta(x))$ which determines the value of loss in the case when the sample has the probability distribution corresponding to hypothesis H_i , but, because of random errors, decision $\delta(x)$ is made.

Making the decision that hypothesis H_i is true in reality true could be one of the hypotheses $H_1, \dots, H_{i-1}, H_{i+1}, \dots, H_s$, i.e, accepting one of the hypothesis, we risk rejecting one of $(S-1)$ really true hypotheses. This risk is called the risk corresponding to the hypothesis H_i , and it is equal to [Berger (1985a), Kachiashvili (2003)]

$$\rho(H_i, \delta) = \int_{\mathbb{R}^n} L(H_i, \delta(x)) p(x/H_i) dx.$$

For any decision rule $\delta(x)$, the complete risk, i.e. the risk of making an incorrect decision, is characterized by the function:

$$r_\delta = \sum_{i=1}^{s} \rho(H_i, \delta(x)) p(H_i) = \sum_{i=1}^{s} p(H_i) \int_{\mathbb{R}^n} L(H_i, \delta(x)) p(x/H_i) dx \tag{1}$$

which is called risk function.

Decision rule $\delta^*(x)$ or, what is the same, Γ_i^*, $i = 1, \dots, s,$- the areas of acceptance of hypotheses H_i, $i = 1, \dots, s,$, are called Bayesian if there takes place:

$$r_{\delta^*} = \min_{\{\delta(x)\}} r_\delta . \tag{2}$$

The Bayesian problem of many hypotheses testing for general and stepwise loss functions has been solved. The obtained decision rules were reduced to concrete working formulae for multivariate normal probability distribution, when the hypotheses are formulated concerning to all parameters of this distribution. For calculation of probability integrals from multivariate normal densities by series using the reduction of dimensionality of multidimensional integrals to one without losing the information were obtained. Formulae for calculation of product moments of normalized normally distributed random variables were also obtained. The problems of existence and continuity of the probability distribution law of linear combination of exponents of quadratic forms of the normally distributed random vector, and also, the problem of finding the closed form of this law were considered. The existence of this law and the opportunity of its unambiguous determination by the calculated moments were proved. The calculation for numerical examples was realized.

The results of research and calculations of concrete examples allow us to infer that, for the Bayesian problem of many hypotheses testing concerning the parameters of multivariate normal distribution, for obtaining correct decisions with high authenticity, the correct choice of loss function depending on the information divergence among the hypotheses is of great importance, and, the correct choice of a priori probabilities of the hypotheses informationally close to the true hypothesis is also of significance.

References

[1] Kachiashvili, K.J. (1989). Bayesian algorithms of many hypothesis testing. Ganatleba, Tbilisi.

[2] Kachiashvili K. J. (2003). *Generalization of Bayesian Rule of Many Simple Hypotheses Testing.* International Journal of Information Technology Decision Making. Vol. 2, No. 1, pp. 41-70.

[3] Berger J.O., B. Boukai, Y., Wang J. (1997). *Unified Frequentist and Bayesian Testing of a Precise Hypothesis.* Statistical Science, Vol. 12, No. 3, pp. 133148

[4] Bernode J.M. (1980). *A Bayesian Analysis of Classical Hypothesis Testing.* Trabajos de Estadística y de Investigación Operativa, Vol. 31, No. 1, 605647.

[5] Cox D. R. (1958). *Some Problems connected with Statistical Inference.* The Annals of Mathematical Statistics. Vol. 29. No. 2 pp.357-372.

[6] Severini, T. A (1994). *Approximately Bayesian Inference.* J. Amer. Statist. Assoc, Vol. 89, No. 425, pp. 242- 249.

[7] Stuart A, Ord K, Arnold S. *Classical Inference and Linear Model.* Vol. 2A., 6th Ed. Oxford University Press Inc.

[8] Weiss, R. (1997). *Bayesian Sample Size Calculations for Hypothesis Testing.* Journal of the Royal Statistical Society. Series D, Vol. 46, No. 2, pp. 185-191.

[9] Shaffer, Juliet Popper (1995) *Multiple hypothesis testing.* Annual Review of Psychology; 46, Health and Medical Complete, pg. 561

[10] James O. Berger (1985) *Statistical Decision Theory and Bayesian Analysis.*

[11] Karl-Rudolf Koch (2007) *Introduction to Bayesian Statistics.* Springer Berlin Heidelberg New York.

[12] Yosef Hochberg *et al*. (1987) *Multiple Comparison Procedures.* John Wiley and Sons.

[13] Laura Ventura, Walter Racugno (2011)Recent advances on Bayesian inference for $P(X<Y)$. Bayesian Analysis, 6, Number 3, pp. 411-428.

[14] Guosheng Yin, (2009) Bayesian Generalized Method of Moments, Bayesian Analysis, 4, Number 2, pp. 191-208.

[15] Andrew Gelman, (2008) Objections to Bayesian statistics, Bayesian Analysis, 3, Number 3, pp. 445-450.

[16] Bruno Lecoutre etl, (2001) Uses, Abuses and Misuses of Significance Tests in the Community: Won't the Bayesian Choice be Unavoidable? International Statistical Review,Volume 69, Issue 3, pages 399-417.

[17] Jos M. Bernardo, Ral Rueda (2002) Bayesian Hypothesis Testing: a Reference Approach, International Statistical Review, Volume 70, Issue 3, pages 351372.